JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 109, 564–590 (1985)

Singular Values and Condition Numbers of Galerkin Matrices Arising from Linear Integral Equations of the First Kind

RICHARD C. ALLEN, JR.

University of New Mexico, Albuquerque, New Mexico 87131

AND

W. ROBERT BOLAND, VANCE FABER, AND G. MILTON WING

Los Alamos National Laboratory, Los Alamos, New Mexico 87545

1. INTRODUCTION

In [1] the problem of the conditioning of matrices arising in the numerical solution of integral equations of the first kind by Galerkin and collocation schemes was investigated. A lower bound on the matrix condition number was found, and its behavior as a function of the smoothness of the kernel of the original equation was studied. Some numerical results presented in [1] and much more extensive studies reported in [2] served to demonstrate the validity and usefulness of the theory.

In these studies all basis functions were required to be orthonormal. Current computational methods frequently use non-orthonormal bases (for example, splines), and the question naturally arises as to whether such bases yield better conditioning. Actually, in both [1] and [2] piecewise constant splines were examined as a special orthonormal set. They, of course, conformed to the general analytical theory. Computationally they usually proved more satisfactory than more classical orthonormal bases, providing both better conditioning and more satisfactory numerical results.

In this paper we shall examine much more general bases including some that are "spline-like." We show that in general little change is made in the analytical results found in [1]. We also provide several corroborating numerical examples.

In [1] it was shown that collocation can be considered as a limiting form of the Galerkin approach. For that reason we confine our entire

0022-247X/85 $3.00

investigation here to the Galerkin scheme. It must be noted, however, that this device is very time consuming as far as actual computations are concerned. Collocation is much more efficient but introduces new problems, such as suitable location of the collocation points, a matter currently being studied.

Because matrix condition numbers are defined in terms of singular values we first present some results relating the singular values of Galerkin matrices to the singular values of the corresponding integral operator. Next condition numbers per se are investigated, both for general bases and for spline-like bases. Finally, numerical studies are presented. Several interesting results are included in the appendixes. They are somewhat peripheral to the principal thrust of the paper.

2. The Problem and Some Notation

We consider the integral equation of the first kind

$$g(x) = \int_0^1 K(x, y)\, f(y)\, dy, \tag{2.1}$$

where $K \in L_2$. Here g is assumed given and f is to be determined. All functions are supposed real. It will often be convenient to write (2.1) in the operator form

$$g = Kf. \tag{2.2}$$

Throughout our work the set $\{\psi_i(x)\}$, $i = 1, 2, \ldots, l$, will be *normal* and linearly independent, but otherwise arbitrary unless further specified. The ψ_j are a basis for some l-dimensional subspace of L_2. We project the solution function f of (2.1) onto the subspace and write approximately

$$f(y) \cong \sum_{j=1}^{l} f_j \psi_j(y). \tag{2.3}$$

Similarly, the set $\{\rho_j(x)\}$ is normal, linearly independent and spans a (probably different) subspace, also assumed l dimensional. The essence of the Galerkin method is to write, using very standard inner product notation,

$$(g, \rho_i) = \sum_{j=1}^{l} f_j (K\psi_j, \rho_i), \qquad i = 1, 2, \ldots, l \tag{2.4}$$

or

$$\mathbf{g} = \mathbf{K}\mathbf{f}, \tag{2.5}$$

where $\mathbf{g}$ and $\mathbf{f}$ are l-vectors, and $\mathbf{K}$ is the matrix whose elements are

$$(K_{ij}) = (K\psi_j, \rho_i). \tag{2.6}$$

At times it will be desirable to study several different bases $\{\psi_i\}$ and $\{\rho_i\}$. Thus $\{\psi_j^{(1)}\}$ may span a space $S^{(1)}$ of dimension $l^{(1)}$, whereas $\{\psi_j^{(2)}\}$ spans $S^{(2)}$ of dimension $l^{(2)}$. Any relationships among the ψ's, the S's, and the l's will be specified as needed. Similar comments hold for the ρ_i. This matter is clarified in Section 3.

Primary interest lies in $\gamma_l(\mathbf{K})$, the condition number of $\mathbf{K}$, defined as

$$\gamma_l(\mathbf{K}) = \sigma_1/\sigma_l, \tag{2.7}$$

where σ_1 and σ_l denote the largest and smallest singular values of $\mathbf{K}$. Results about γ_l will often be obtained in terms of the behavior of μ_l, the lth singular value of the integral operator K. As in [1] we shall be interested in obtaining a lower bound on γ_l in terms of the behavior of $K(x, y)$.

3. Some General Results Concerning Bases and Singular Values

We begin by noting that the domain of an L_2 kernel $K(x, y)$ is all of L_2, but the range may well be a proper subset W of L_2. Often it is difficult to determine W without deep analysis of the operator K. This can have important consequences for the Galerkin method because clearly one wishes the functions ρ_j to do a reasonable job of spanning W; see (2.4).

To understand these matters somewhat better, let V_l and W_l be two sequences of subspaces of L_2. (Note that W_l may not be contained in W for given l.) We require V_l and W_l to have the following properties:

A. $\dim V_l = \dim W_l = l$.

B. V_l is generated by the *normal* basis $\Psi_l = \{\psi_j^l\}$, $j = 1, 2, \ldots, l$.
   W_l is generated by the *normal* basis $\mathrm{P}_l = \{\rho_j^l\}$, $j = 1, 2, \ldots, l$.

We define the matrix of K relative to the bases Ψ_l and P_l by

$$\mathbf{K}^l = (K_{ij}^l) = (K\psi_j^l, \rho_i^l). \tag{3.1}$$

Note that (3.1) agrees with (2.6) except that the order l of $\mathbf{K}$ is made explicit. We observe also that in practice the sets Ψ_l and P_l are usually first picked and the spaces V_l and W_l then formed.

Recall that nth singular value σ_n^l of $\mathbf{K}^l$, $n \leqslant l$, is given by

$$(\sigma_n^l)^2 = \max_{S_n \subseteq R^l} \min_{\substack{\mathbf{u} \in S_n \\ \|\mathbf{u}\|_n = 1}} \|\mathbf{Ku}\|_n, \tag{3.2}$$

where R^l is the Euclidean vector space of dimension l and S_n is any n-dimensional subspace of R^l. We use the notation

$$\|\mathbf{u}\|_n^2 = \sum_{j=1}^{n} u_j^2. \tag{3.3}$$

We write (3.2)

$$(\sigma_n^l)^2 = \max_{S_n \subseteq R^l} \min_{\substack{\mathbf{u} \in S_n \\ \|\mathbf{u}\|_n = 1}} \sum_{i=1}^{l} \left[\sum_{j=1}^{l} u_j (K\psi_j^l, \rho_i^l) \right]^2. \tag{3.4}$$

If we assume Ψ_l and P_l are orthogonal as well as normal Bessel's inequality and (3.4) yield

$$(\sigma_n^l)^2 \leqslant \max_{A_n \subseteq V_l} \min_{\substack{v \in A_n \\ \|v\| = 1}} \|Kv\|^2 = (\mu_n^l)^2, \tag{3.5}$$

where A_n is any n-dimensional subspace of V_l. Here $\| \cdots \|$denotes the ordinary L_2 norm. The right side of (3.5) is just the nth singular value of K restricted to the domain V_l. Clearly

$$\max_{A_n \subseteq V_l} \min_{\substack{V \in A_n \\ \|v\| = 1}} \|Kv\|^2 \leqslant \max_{A_n \subseteq L_2} \min_{\substack{V \in A_n \\ \|v\| = 1}} \|Kv\|^2 = \mu_n^2, \tag{3.6}$$

where μ_n is the nth singular value of K.

THEOREM 3.1. *If the bases Ψ_l and P_l are orthonormal, then*

$$\sigma_n^l \leqslant \mu_n^l \leqslant \mu_n, \qquad n = 1, 2, \ldots, l. \tag{3.7}$$

This theorem strengthens a result of [1] where the inequality $\sigma_l^l \leqslant \mu_l$ was established. It also generalizes some results of J. Hersch [3] (see Appendix A).

We now drop the assumption of orthogonality of the bases and return to (3.4). Recall that W_l is not assumed to be included in W. Let Π_l be the operator that projects W_l on $K(V_l)$, the range of K restricted to V_l. Clearly (3.4) can be rewritten

$$(\sigma_n^l)^2 = \max_{S_n \subseteq R^l} \min_{\substack{\mathbf{u} \in S_n \\ \|\mathbf{u}\|_n = 1}} \sum_{i=1}^{l} \left[\sum_{j=1}^{l} u_j (K\psi_j^l, \pi_l \rho_i^l) \right]^2. \tag{3.8}$$

If $\pi_l \rho_i^l$ is the null vector for all i, then σ_n^l is obviously zero, and the corresponding Galerkin matrix is singular and of no interest computationally. We also note that if $\dim K(V_l) < l$, K^l will also be singular. We assume henceforth that neither event occurs. Obviously in practice one must use some care in selecting basis functions.

Next choose an orthonormal basis for V_l, denoted Θ_l, and one for $K(V_l)$, denoted Ξ_l. There exists a linear transformation P^l from V_l onto V_l such that

$$\psi_j^l = P^l \xi_j^l, \qquad j = 1, 2,..., l. \tag{3.9}$$

Similarly there exists a Q^l from $K(V_l)$ onto $K(V_l)$ such that

$$\pi_l \rho_i^l = Q^l \theta_i^l, \qquad i = 1, 2,..., l. \tag{3.10}$$

Using (3.8) we obtain

$$\begin{aligned}
(\sigma_n^l)^2 &= \max_{S_n \subseteq R^l} \min_{\substack{\mathbf{u} \in S_n \\ \|\mathbf{u}\|_n = 1}} \sum_{i=1}^{l} \left[\sum_{j=1}^{l} u_j (KP^l \xi_j^l, Q^l \theta_i^l) \right]^2 \\
&= \max_{S_n \subseteq R^l} \min_{\substack{\mathbf{u} \in S_n \\ \|\mathbf{u}\|_n = 1}} \sum_{i=1}^{l} \left[\sum_{j=1}^{l} u_j (Q^{l*} KP^l \xi_j^l, \theta_i^l) \right]^2 \\
&= \max_{A_n \subseteq V_l} \min_{\substack{v \in A_n \\ \|v\| = 1}} \|Q^{l*} KP^l v\|^2.
\end{aligned} \tag{3.11}$$

In this representation the σ_n^l are clearly the singular values of K (restricted to V^l) in the bases Θ_l and Ξ_l.

We shall find it convenient to define

$$(\lambda_n^l)^2 = \max_{A_n \subseteq V_l} \min_{\substack{v \in A_n \\ \|v\| = 1}} \|Q^{l*} Kv\|^2. \tag{3.12}$$

A bit of calculation reveals that the λ_n^l are the singular values of K (restricted to V_l) in the bases Ψ_l and Φ_l.

We can now prove

THEOREM 3.2. *For all $n \leqslant l$, and for Ψ_l only a normal basis,*

$$\frac{1}{(\lambda_n^l)^2} \geqslant \frac{1}{l} \sum_{j=1}^{n} \frac{1}{(\mu_j^l)^2} \geqslant \frac{1}{l} \sum_{j=1}^{n} \left(\frac{1}{\mu_j}\right)^2. \tag{3.13}$$

Remark. Our assumptions prevent any of the denominators in (3.13) from being zero. These assumptions may be relaxed, in which case (3.13) continues to hold in the event of a zero denominator, with proper interpretation.

Proof. We construct the trace of $Q^{l*}Q^l$:

$$\begin{aligned}\operatorname{Tr}(Q^{l*}Q^l) &= \sum_{j=1}^{l} (Q^{l*}Q^l\theta_j^l, \theta_j^l)\\ &= \sum_{j=1}^{l} (\pi_l\rho_j^l, \pi_l\rho_j^l)\\ &\leqslant \sum_{j=1}^{l} \|\rho_j^l\|^2 = l. \end{aligned}\tag{3.14}$$

In (3.12) let $\bar{A}_n$ be the subspace of V_l on which the maximum is actually achieved, and consider K restricted to $\bar{A}_n$. The singular values of this operator are given by

$$\alpha_j^n = \max_{A_j \subset \bar{A}_n} \min_{\substack{u \in A_j\\ \|u\|=1}} \|Ku\|, \qquad j = 1, 2, \ldots, n. \tag{3.15}$$

Further, let (v_j, w_j), $j = 1, 2, \ldots, n$, be the left and right normalized singular vectors of this operator,

$$Kv_j = \alpha_j^n w_j. \tag{3.16}$$

From (3.12)

$$\begin{aligned}(\lambda_n^l)^2 &= \min_{\substack{v \in \bar{A}_n\\ \|v\|=1}} \|Q^{l*}Kv\|^2\\ &\leqslant \|Q^{l*}Kv_j\|^2\\ &\leqslant (\alpha_j^n)^2 \|Q^{l*}w_j\|^2, \qquad j = 1, 2, \ldots, n. \end{aligned}\tag{3.17}$$

Now from (3.5) and (3.15)

$$(\alpha_j^n)^2 \leqslant (\mu_j^l)^2, \tag{3.18}$$

so that

$$(\lambda_n^l)^2 \leqslant (\mu_j^l)^2 \|Q^{l*}w_j\|^2, \qquad j = 1, 2, \ldots, n. \tag{3.19}$$

Thus

$$\sum_{j=1}^{l} \frac{(\lambda_n^l)^2}{(\mu_j^l)^2} \leqslant \sum_{j=1}^{l} (Q^l Q^{l*} w_j, w_j)$$
$$= \operatorname{Tr}(Q^l Q^{l*}), \tag{3.20}$$

since the trace is invariant under orthogonal transformation. The first part of the inequality (3.13) now follows from (3.20) and (3.14). The second part is obtained from (3.5) and (3.6).

We obtain one more result of this kind; it will prove very useful in Section 4:

THEOREM 3.3. *For Ψ_l and R_l only normal bases*

$$\frac{1}{(\sigma_l^l)^2} \geqslant \frac{1}{l^2} \sum_{j=1}^{l} (l-j+1) \frac{1}{(\mu_j^l)^2}$$
$$\geqslant \frac{1}{l^2} \sum_{j=1}^{l} (l-j+1) \frac{1}{(\mu_j)^2}. \tag{3.21}$$

Proof. From (3.11):

$$(\sigma_l^l)^2 = \min_{\substack{v \in V_l \\ \|v\| = 1}} \|Q^{l*} K P^l v\|^2$$
$$= \min_{\substack{v \in K(V_l) \\ \|v\| = 1}} \|P^{l*} K^* Q^l v\|^2, \tag{3.22}$$

since an operator and its conjugate have the same singular values. We now repeat the argument of Theorem 3.2, replacing Q^l by P^l and K by $(Q^{l*}K)^*$, to obtain at once

$$\frac{1}{(\sigma_l^l)^2} \geqslant \frac{1}{l} \sum_{j=1}^{l} \frac{1}{(\lambda_j^l)^2}. \tag{3.23}$$

Applying (3.13) gives

$$\frac{1}{(\sigma_l^l)^2} \geqslant \frac{1}{l} \sum_{j=1}^{l} \left[\frac{1}{l} \sum_{k=1}^{j} \frac{1}{(\mu_k^l)^2} \right] \tag{3.24}$$
$$= \frac{1}{l^2} \sum_{j=1}^{l} (l-j+1) \frac{1}{(\mu_j^l)^2}. \tag{3.25}$$

The remaining part of (3.21) follows from (3.5) and (3.6).

COROLLARY 3.4. *For Ψ_l and P_l merely normal*

$$\sigma_l^l \leqslant l\mu_l. \tag{3.26}$$

Proof. We simply observe in (3.21) that

$$\sum_{j=1}^{l} (l-j+1)\frac{1}{(\mu_j)^2} \geqslant \frac{1}{(\mu_l)^2}. \tag{3.27}$$

We turn now to a discussion of σ_1^l, of interest because of the definition of the condition number (see (2.7)). To obtain a lower bound on this number, it is desirable to find a lower bound on σ_1^l. Unfortunately, this seems to be very hard to accomplish except in very special circumstances. For a special class of non-orthogonal basis functions we shall examine the matter in Section 5. That will actually generalize a result of [1]. We also call attention to the result on orthonormal bases found in [1].

Some generalization of the latter is possible in our current framework, provided we require that the subspaces of V_l "fill out" L_2. More precisely, we ask that for $\varepsilon > 0$ and $v \in L_2$ there exists M such that for $l > M$ one can find $v_l \in V_l$ such that $\|v_l - v\| < \varepsilon$. Note that there is no requirement that the V_l be nested. The sequence V_l is, in a sense, "closed in L_2."

THEOREM 3.5. *Let the sequence of subspaces V_l be closed in L_2 in the sense above. Let Ψ_l and P_l be* orthonormal *bases in V_l and W_l. Then*

$$\lim_{l \to \infty} \sigma_1^l = \mu_1. \tag{3.28}$$

Proof. Recall from Theorem 3.1 that for all l

$$\sigma_1^l \leqslant \mu_1. \tag{3.29}$$

Choose $0 < \varepsilon < \mu_1$. Then there exists M such that for $l > M$ there is a $v_l \in V_l$ such that

$$\|v_l - u_1\| < \frac{\varepsilon}{2\mu_1}, \tag{3.30}$$

where u_1 is the first normalized right singular vector of K. Thus

$$\begin{aligned} \|v_l\| &\leqslant \|u_1\| + \|v_l - u_1\| \\ &\leqslant 1 + \frac{\varepsilon}{2(\mu_1 - \varepsilon)} = \frac{2\mu_1 - \varepsilon}{2(\mu_1 - \varepsilon)}. \end{aligned} \tag{3.31}$$

Moreover, by (3.5),

$$\|K(v_l - u_1)\| \leqslant \mu_1 \|v_l - u_1\|, \tag{3.32}$$

from which it follows that

$$\|Ku_1\| \leqslant \|Kv_l\| + \frac{\varepsilon}{2}, \tag{3.33}$$

and so

$$\frac{\|Kv_l\|}{\|v_l\|} \geqslant \frac{\mu_1 - \varepsilon/2}{(2\mu_1 - \varepsilon)} 2(\mu_1 - \varepsilon) = \mu_1 - \varepsilon. \tag{3.34}$$

Recall that since Ψ_l and R_l are orthonormal

$$\sigma_1^l = \max_{\substack{v \in V_l \\ v \neq 0}} \frac{\|Kv\|}{\|v\|} \geqslant \frac{\|Kv_l\|}{\|v_l\|} = \mu_1 - \varepsilon. \tag{3.35}$$

Using (3.29) gives

$$\mu_1 - \varepsilon \leqslant \sigma_1^l \leqslant \mu_1, \tag{3.36}$$

and the desired result.

We observe that the result of [1] concerning σ_1 for orthonormal systems follows at once if the system $\{\phi_j\}$ of [1] is complete in L_2.

4. Applications of the Results of Section 3 to Integral Equations with Smooth Kernels

Numerous results are available relating the behavior of singular values of the integral operator K to the smoothness of $K(x, y)$ (see [1]). We state just two of these:

THEOREM 4.1 (Chang [4]). *Let $K(x, y)$ be in L_2 on $[0, 1] \times [0, 1]$. Suppose*

(a) $K, \dfrac{\partial K}{\partial x}, \dfrac{\partial^2 K}{\partial x^2}, \ldots, \dfrac{\partial^{s-1} K}{\partial x^{s-1}}$

exist and are continuous in x for almost all y and

(b) $\dfrac{\partial^s K(x, y)}{\partial x^s} = \displaystyle\int_0^x g(t, y)\, dt + A(y),$

where $g \in L_2$ *and* $A(y)$ *is integrable. Then*

$$\mu_l = \varepsilon_l l^{-s-3/2}, \tag{4.1}$$

where $\lim_{l \to \infty} \varepsilon_l = 0$.

THEOREM 4.2 (Hille and Tamarkin [5]). *Let* $K(x, y)$ *be in* L_2 *and* symmetric *and satisfy* (a) *of Theorem* 4.1. *Further suppose that* g *in* (b) *of that theorem is such that instead of* $g \in L_2$ *we have just*

$$\int_0^1 \left[\int_0^1 |g(x, y)|^p \, dx \right]^{1/(p-1)} dy < \infty$$

for some p, $1 < p \leqslant 2$. *Then* (4.1) *can be replaced by*

$$\mu_l = \varepsilon_l l^{-s-2+1/p}. \tag{4.2}$$

In each of these theorems we have

$$\mu_l < M l^{-\alpha}, \qquad l = 1, 2, \ldots, \tag{4.3}$$

for some $M > 0$ and appropriate α. From Theorem 3.3 we obtain

$$\begin{aligned} \frac{1}{(\sigma_l^l)^2} &\geqslant \frac{1}{M^2 l^2} \sum_{j=1}^{l} (l-j+1)\, l^{2\alpha} \\ &\geqslant M^l l^{2\alpha}. \end{aligned} \tag{4.4}$$

This yields

THEOREM 4.3. *If the hypotheses of Theorem* 4.1 *hold, then*

$$\sigma_l^l \leqslant M'' l^{-s-3/2}. \tag{4.5}$$

Similarly, the hypotheses of Theorem 4.2 *yield*

$$\sigma_l^l \leqslant M'' l^{-s-2+1/p}. \tag{4.6}$$

This result holds for completely arbitrary normal bases Ψ_l *and* P_l *It differs from that in* [1] *for orthonormal bases only in that in* [1] M'' *in* (4.5) *and* (4.6) *is replaced by* ε_l' *where* $\varepsilon_l' \to 0$.

This slightly sharper result of [1] is obtained through the use of Theorem 3.1 of this paper. For arbitrary bases we have only (3.21). However, this may be improved by considering spline-like bases, and there is then obtained the analogues of (4.5) and (4.6) with M'' replaced by ε_l'. For our analysis of condition numbers, Theorem 4.3 usually provides adequate information. Therefore the improvement obtained for spline-like Ψ_l and P_l is consigned to Appendix B.

We observe at once

THEOREM 4.4. *If $\sigma_1^l \geqslant m > 0$ for $l = 1, 2, \ldots$, then the condition number $\gamma_l(\mathbf{K}^l)$, where*

$$\gamma_l(\mathbf{K}^l) = \sigma_1^l / \sigma_l^l, \tag{4.7}$$

satisfies

$$\gamma_l(\mathbf{K}^l) \geqslant M''' l^\alpha, \tag{4.8}$$

where

$$\alpha = s + \tfrac{3}{2} \qquad (\alpha = s + 2 - 1/p)$$

provided the assumptions of Theorem 4.1 (Theorem 4.2) are satisfied.

5. A CONSIDERATION OF σ_1^l FOR BASES OF SMALL SUPPORT

As observed in Section 3, the analysis of σ_1^l is surprisingly complicated. It can be shown that for certain basis sets

$$\lim_{l \to \infty} \sigma_1^l = 0.$$

In fact, for such sets we can have $\gamma_l(\mathbf{K}^l) = 1$ for all l. We prefer to leave such pathological considerations for Appendix C, and turn to cases that are more likely to arise in practice.

Numerous experimental calculations suggest that, in general, Galerkin matrices generated by basis functions of small support (e.g., B-splines) are better conditioned than those arising from more classical bases with global support (e.g., Legendre and Tchebychev polynomials). We therefore study such Ψ_l and P_l.

Choose a doubly infinite sequence of points $\{x_j^l\}$, where

$$\lim_{j \to -\infty} x_j^l = -\infty,$$

$$\lim_{j \to \infty} x_j^l = +\infty.$$

Require that for some fixed β

$$x_{j+1}^l - x_j^l < \frac{\beta}{l}. \tag{5.1}$$

Next select a sequence of functions $f_j^l(x)$, $j=0, \pm 1, \pm 2, \ldots$, with supports

$$x_i^l \leqslant x \leqslant x_{i+j}^l \qquad j \leqslant k, \tag{5.2}$$

k independent of i and l, and require that each f_i^l be positive on the interior of its support. Let $\chi(x)$ be the characteristic function of $[0, 1]$ and define

$$\hat{f}_i^l(x) = f_i^l(x)\, \chi(x). \tag{5.3}$$

Finally, let $\psi_i^l(x)$ be constructed by defining

$$\psi_i^l(x) = \tilde{f}_i^l(x), \tag{5.4}$$

where $\tilde{f}_i^l$ are the normalized $\hat{f}_i^l$'s. The ψ_i^l's form an acceptable basis set Ψ_l.

Obviously, the set Ψ_l is quite arbitrary. We impose the following condition. Let the support of $\psi_i^l(x)$ be on the interior of $[0, 1]$. Call the length of the support h_i^l. Then there exists a constant $c > 0$, independent of i and l, such that

$$\int_0^1 \psi_i^l(x)\, dx \geqslant c\sqrt{h_i^l}. \tag{5.5}$$

Observe that (5.5) does not follow just from the fact that the ψ_i^l are normal.

Before proceeding, we mention two examples of the kind of basis set just described. The first is a very special case of the second.

EXAMPLE 1. Let $k=1$ in (5.2). Choose $f_i^l(x) \equiv 1$. Suppose the support $[x_i^l, x_{i+1}^l]$ of f_i^l is interior to $[0, 1]$, so $\hat{f}_i^l = f_i^l$. Then

$$\begin{aligned} \psi_i^l(x) &= \frac{1}{\sqrt{x_{i+1}^l - x_i^l}}\, f_i^l(x) \\ &= \frac{1}{\sqrt{h_i^l}}\, f_i^l(x). \end{aligned} \tag{5.6}$$

Thus

$$\begin{aligned} \int_0^1 \psi_i^l(x)\, dx &= \int_{x_i^l}^{x_{i+1}^l} \frac{1}{\sqrt{h_i^l}}\, dx \\ &= \sqrt{h_i^l}. \end{aligned} \tag{5.7}$$

Thus $c=1$ in (5.5).

The ψ_i^l are, of course, just the column functions discussed in [1].

EXAMPLE 2. Fix j in (5.2), and choose any normalized $F(x)$, where F has support $[0, 1]$ and is positive on the interior. Define

$$f_i^l(x) = \frac{1}{\sqrt{h_i^l}} F[(x - x_i)/(h_i^l)]. \tag{5.8}$$

For $0 < x_j < x_{i+j} < 1$, we find

$$\int_0^1 \{f_i^l(x)\}^2 \, dx = \frac{1}{h_i^l} \int_{x_i}^{x_{i+j}} F^2[(x - x_i)/(h_i^l)] \, dx. \tag{5.9}$$

Therefore we have

$$\psi_i^l(x) = \frac{1}{\sqrt{h_i^l}} F[(x - x_i)/(h_i^l)], \qquad 0 < x_i < x_{i+j} < 1 \tag{5.10}$$

and

$$\int_0^l \psi_i^l(x) \, dx = \frac{1}{\sqrt{h_i^l}} \int_{x_i}^{x_{i+j}} F[(x - \mathrm{x}_i)/(h_i^l)] \, dx = \sqrt{h_i^l}. \tag{5.11}$$

Again (5.5) is satisfied with $c = 1$.

As a special case of this example, we cite the cardinal splines of order j, obtained by choosing a particular F and a particular sequence $\{x_i^l\}$ (see [6]). In Example 1 we have $F(x) \equiv 1$, $0 \leqslant x \leqslant 1$, and $j = 1$.

The condition (5.5) is required in the proof of Theorem 5.1. The fact that (5.5) need only hold for ψ_j^l with support interior to $[0, 1]$ is a convenient one since the behavior of basis elements whose construction is a bit awkward as a result of (5.3) can be totally disregarded. We have ignored them in the two examples.

DEFINITION. We shall call any set Ψ_l defined as above a generalized spline basis (GSB).

We now let P_l also be a GSB, with the notations $\tilde{x}_i^l$, $\tilde{h}_i^l$, $\tilde{c}$ used in an obvious way, and proceed to a study of σ_1^l.

Suppose that there is a square S in $[0, 1] \times [0, 1]$ such that

$$K(x, y) \geqslant M > 0, \qquad (x, y) \in S. \tag{5.12}$$

We compute

$$K_{ij}^l = \int_S \int K(x, y) \, \psi_j^l(y) \, \rho_i^l(x) \, dx \, dy$$

$$\geqslant M \int_0^1 \psi_j^l(y) \, dy \int_0^1 \rho_i^l(x) \, dx \tag{5.13}$$

provided the supports of the basis functions are in S. Conditions (5.1) and (5.2) assure that this will be the case for l sufficiently large, provided i lies in some set I and j in J. From (5.13) and (5.5)

$$K_{ij}^l > Mc\tilde{c}\sqrt{h_i^l \tilde{h}_j^l}, \qquad i \in I,\ j \in J. \tag{5.14}$$

Recall that

$$(\sigma_1^l)^2 \geqslant \sum_{i=1}^{l} \left(\sum_{j=1}^{l} K_{ij}^l u_j \right)^2, \qquad \|u\|_l = 1. \tag{5.15}$$

Pick u so that

$$\begin{aligned} u_j &= A\sqrt{\tilde{h}_j^l}, \qquad & j \in J \\ &= 0, & j \notin J, \end{aligned} \tag{5.16}$$

where A is determined by the condition $\|u\|_l = 1$. Thus

$$1 = A^2 \sum_{j \in J} \tilde{h}_j^l \leqslant A^2 L, \tag{5.17}$$

where L is the length of a side of S. Hence

$$A^2 \geqslant \frac{1}{L}. \tag{5.18}$$

Note also that for l sufficiently large

$$\sum_{j \in J} \tilde{h}_j^l \geqslant \frac{L}{2}, \qquad \sum_{i \in I} h_i^l \geqslant \frac{L}{2}. \tag{5.19}$$

Returning to (5.15) we find

$$\begin{aligned} (\sigma_1^l)^2 &\geqslant \sum_{i \in I} \left[\sum_{j \in J} Mc\tilde{c}\sqrt{h_i^l \tilde{h}_j^l}\, A\sqrt{\tilde{h}_j^l} \right]^2 \\ &\geqslant \frac{M^2 c^2 \tilde{c}^2}{L} \sum_{i \in I} h_i^l \left(\sum_{j \in J} \tilde{h}_j^l \right)^2 \\ &\geqslant M^2 c^2 \tilde{c}^2 \frac{L^2}{8} = m > 0. \end{aligned} \tag{5.20}$$

We have proved

THEOREM 5.1. *If the bases Ψ_l and P_l are both* GSB *and if $K(x, y)$ does not vanish on some square, then $\sigma_1^l \geqslant m > 0$ for some m and all l sufficiently large. Thus Theorem* 4.4 *applies.*

6. Some Numerical Examples

In this section we tabulate the condition numbers of the $\mathbf{K}^l$ matrix for the Galerkin method for two different kernels $K(x, y)$ using ordinary B-splines as basis functions for both the $f(x)$ and $g(x)$ expansions. Our results are not intended to be comprehensive, but rather illustrative.

Actually many kernels were considered during the course of our investigations, but we present results only for the following:

$$K(x, y = e^{-\alpha|x-y|}; \qquad -1 \leqslant x, y \leqslant 1; \alpha = 0.001, 0.1, 1.0, 10.0 \quad (6.1)$$

$$K(x, y) = |x-y|^{\beta}; \qquad -1 \leqslant x, y \leqslant 1; \beta = 2.0, 2.001, 2.1, 2.5. \quad (6.2)$$

For the first kernel, K is continuous but has a discontinuous first derivative. For the second, K is smoother, having a continuous second derivative but not a third, except when $\beta = 2.0$ for which K possesses all derivatives. From the theory we would expect these differences in smoothness to be reflected in the magnitude of the computed condition numbers.

In the numerical results that follow all integrals and related quantities were approximated using the B-spline routines available in the SLATEC library [7]. All matrix condition numbers were estimated using the LINPACK routine SGECO [8]. Programs were written in Fortran and computations were carried out in single precision on the Cray-1 computers at the Los Alamos National Laboratory.

To examine the behavior of the condition number $\gamma(\mathbf{K}^l)$ as a function of l, we tabulated γ for $l = 10, 20, 30, 40, 50$ and then fitted the results to the function

$$\gamma(\mathbf{K}^l) = cl^{\eta}$$

in the sense of least squares. The form of this function was, of course, suggested by the theory. Results predicted by the theory are given for comparison in the form cl^{η} where c may depend on η.

Tables I–VII follow, the first four giving conditions numbers $\gamma(\mathbf{K}^l)$ for Galerkin matrices for the kernel $K(x, y) = e^{-\alpha|x-y|}$ and the last three for $K(x, y) = |x-y|^{\beta}$. The case $\beta = 2.0$ is omitted since all of the condition numbers were of the order of 10^{16}. This should be expected, of course, since $|x-y|^{2.0}$ is infinitely differentiable. In all tables the notation $0 \cdot pq(n)$ means $0 \cdot pq \times 10^n$ where p and q are decimal integers.

It is not surprising that $\gamma(\mathbf{K}^l)$ increases with l. It should also be observed that in many cases γ increases with k, the order of the B-spline. The least squares fit is remarkably close to the lower bound predicted by the theory and suggests that the numerical schemes employed are very sensitive to the behavior of relatively high order derivatives of K.

TABLE I

$K = e^{-0.001|x-y|}$ B-Spline Bases of Order k

k / l	1	2	3	4
10	0.27(6)	0.69(6)	0.25(7)	0.76(7)
20	0.11(7)	0.29(7)	0.10(8)	0.35(8)
30	0.26(7)	0.66(7)	0.22(8)	0.76(8)
40	0.47(7)	0.12(8)	0.38(8)	0.13(9)
50	0.74(7)	0.18(8)	0.58(8)	0.19(9)
Fit	0.24(4) $l^{2.06}$	0.62(4) $l^{2.04}$	0.28(5) $l^{1.95}$	0.78(5) $l^{2.01}$
Theory	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$

TABLE II

$K = e^{-0.1|x-y|}$ B-Spline Bases of Order k

k / l	1	2	3	4
10	0.26(4)	0.66(4)	0.24(5)	0.73(5)
20	0.11(5)	0.28(5)	0.97(5)	0.33(6)
30	0.25(5)	0.63(5)	0.21(6)	0.73(6)
40	0.45(5)	0.11(6)	0.36(6)	0.12(7)
50	0.70(5)	0.18(6)	0.54(6)	0.18(7)
Fit	0.23(4) $l^{2.06}$	0.60(2) $l^{2.04}$	0.27(3) $l^{1.95}$	0.74(3) $l^{2.01}$
Theory	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$

TABLE III

$K = e^{-1.0|x-y|}$ B-Spline Bases of Order k

k / l	1	2	3	4
10	0.17(3)	0.45(3)	0.16(4)	0.49(4)
20	0.72(3)	0.18(4)	0.65(4)	0.22(5)
30	0.17(4)	0.42(4)	0.14(5)	0.43(5)
40	0.30(4)	0.75(4)	0.24(5)	0.82(5)
50	0.47(4)	0.12(5)	0.36(5)	0.12(6)
Fit	0.15(1) $l^{2.06}$	0.42(1) $l^{2.03}$	0.19(2) $l^{1.94}$	0.50(2) $l^{2.01}$
Theory	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$

TABLE IV

$K = e^{-10.0|x-y|}$ B-Spline Bases of Order k

k / l	1	2	3	4
10	0.39(1)	0.12(2)	0.41(2)	0.12(3)
20	0.13(2)	0.34(2)	0.12(3)	0.39(3)
30	0.27(2)	0.72(2)	0.24(3)	0.82(3)
40	0.48(2)	0.12(3)	0.40(3)	0.14(4)
50	0.75(2)	0.19(3)	0.60(3)	0.20(4)
Fit	0.55(−1) $l^{1.83}$	0.22(0) $l^{1.66}$	0.88(0) $l^{1.66}$	0.21(1) $l^{1.75}$
Theory	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$	$cl^{1.5}$

TABLE V

$K = |x-y|^{2.001}$ B-Spline Bases of Order k

k / l	1	2	3	4
10	0.15(7)	0.19(7)	0.74(7)	0.19(8)
20	0.13(8)	0.29(8)	0.10(9)	0.34(9)
30	0.47(8)	0.11(9)	0.43(9)	0.16(10)
40	0.12(9)	0.27(9)	0.11(10)	0.43(10)
50	0.23(9)	0.55(9)	0.22(10)	0.91(10)
Fit	0.11(4) $l^{3.15}$	0.65(3) $l^{3.52}$	0.21(4) $l^{3.57}$	0.30(4) $l^{3.84}$
Theory	$cl^{3.001}$	$cl^{3.001}$	$cl^{3.001}$	$cl^{3.001}$

TABLE VI

$K = |x-y|^{2.1}$ B-Spline Bases of Order k

k / l	1	2	3	4
10	0.13(5)	0.31(5)	0.87(5)	0.23(6)
20	0.17(6)	0.38(6)	0.14(7)	0.45(7)
30	0.65(6)	0.15(7)	0.58(7)	0.22(8)
40	0.16(7)	0.38(7)	0.16(8)	0.61(8)
50	0.34(7)	0.80(7)	0.32(8)	0.13(9)
Fit	0.10(2) $l^{3.25}$	0.11(2) $l^{3.45}$	0.20(2) $l^{3.63}$	0.28(2) $l^{3.95}$
Theory	$cl^{3.1}$	$cl^{3.1}$	$cl^{3.1}$	$cl^{3.1}$

TABLE VII

$K = |x-y|^{2.5}$ B-Spline Bases of Order k

k / l	1	2	3	4
10	0.64(4)	0.13(5)	0.36(5)	0.95(5)
20	0.11(6)	0.24(6)	0.76(6)	0.27(7)
30	0.50(6)	0.11(7)	0.42(7)	0.16(8)
40	0.14(7)	0.33(7)	0.13(8)	0.52(8)
50	0.33(7)	0.75(7)	0.28(8)	0.13(9)
Fit	0.94(0) $l^{3.85}$	0.16(1) $l^{3.93}$	0.27(1) $l^{4.16}$	0.35(1) $l^{4.48}$
Theory	$cl^{3.5}$	$cl^{3.5}$	$cl^{3.5}$	$cl^{3.5}$

7. Some Remarks and Conclusions

We have shown here that the lower bounds for condition numbers of Galerkin matrices found in [1] under highly restrictive assumptions on the bases carry over virtually unchanged to a much wider class of general normal bases. Somewhat special attention has been given to spline-like bases, partly because of their frequent use in computations. Numerical calculations presented indicate that, just as in [1, 2], the lower bounds seem remarkable good.

Our results are again of a negative kind. Certainly, results on upper bounds of condition numbers are very desirable. Although the numerical computations suggest that the lower bounds may often be quite close to the upper bounds, the proof eludes us.

APPENDIX A: Connection with Some Results of J. Hersch

We have shown in Theorem 3.1 that for Ψ_l and P_l orthonormal we have

$$\sigma_n^l \geqslant \mu_n, \qquad n = 1, 2, ..., l. \tag{A1}$$

It follows that if $\Psi_l = \mathrm{P}_l$ and if K is symmetric and positive definite, then

$$\mathrm{Tr}(\mathbf{K}^l) = \sum_{n=1}^{l} \sigma_n^l \leqslant \sum_{n=1}^{l} \mu_n. \tag{A2}$$

If we choose the basis functions to be the eigenfunctions of K, then equality holds in (A2). This observation leads to

$$\sum_{n=1}^{l} \mu_n = \max_{A_l \subset L_2} \operatorname{Tr}(\mathbf{K}^l). \tag{A3}$$

Applying the same kind of reasoning to the inverse matrix $(\mathbf{K}^l)^{-1}$, we find

$$\sum_{n=1}^{l} \mu_n^{-1} = \min_{A_l \subset L_2} \operatorname{Tr}((\mathbf{K}^l)^{-1}). \tag{A4}$$

These results, and many others more general, have been established by J. Hersch [3] (see also [9]).

APPENDIX B: A Further Consideration of σ_l^l and Its Implications in the Case of GSB Bases

We recall from (3.8) that

$$(\sigma_l^l)^2 = \min_{\|\mathbf{u}\|_l = 1} \sum_{i=1}^{l} \left[\sum_{j=1}^{l} u_j (K\psi_j^l, \pi_l \rho_i^l) \right]^2 \tag{B1}$$

and consider Ψ_l and P_l merely normal. Define

$$U^l = \sum_{j=1}^{l} u_j \psi_j^l. \tag{B2}$$

Unfortunately, $\|U^l\| \neq 1$, in general, despite the constraint $\|\mathbf{u}\|_l = 1$. Rather

$$\|U^l\|^2 \leqslant \Lambda, \tag{B3}$$

where Λ is the maximum eigenvalue of the Grammian matrix with elements (ψ_i^l, ψ_j^l). It is easy to verify that

$$\begin{aligned} (\sigma_l^l)^2 &\leqslant \min_{\|U^l\| = \Lambda^{1/2}} \sum_{i=1}^{l} (KU^l, \pi_l \rho_i^l)^2 \\ &= \min_{\|U^l\| = 1} \Lambda \sum_{i=1}^{l} (KU^l, \pi_l \rho_i^l)^2. \end{aligned} \tag{B4}$$

Now let Θ_l be an orthonormal basis for $K(V_l)$ as in Section 3. Suppose $\tilde{W}^l$, where $\|\tilde{W}^l\| = 1$ provides the minimum in (B4). Write

$$K\tilde{W}^l = \sum_{j=1}^{l} w_j \theta_j^l \tag{B5}$$

and, expanding the notation (3.10) somewhat,

$$\pi_l \rho_i^l = \sum_{k=1}^{l} q_{ik}^l \theta_k^l. \tag{B6}$$

Thus

$$(K\tilde{W}^l, \pi^l \rho_i^l)_l = \sum_{j=1}^{l} q_{ij}^l w_j, \tag{B7}$$

and (B4) yields

$$\begin{aligned}(\sigma_l^l)^2 &\leqslant \Lambda \sum_{i=1}^{l} \left(\sum_{j=1}^{l} q_{ij}^l w_j \right)^2 \\ &= \Lambda \, \|\mathbf{Q}^l \mathbf{w}\|_l^2. \end{aligned} \tag{B8}$$

Note that if $\|\mathbf{w}\|_l^2$ were equal to 1, then the maximum value of the right side of (B8) would be $\Lambda \Gamma^2$, where Γ is the largest singular value of $\mathbf{Q}^l$. Since by (B5) we have

$$\|\mathbf{w}\|_l^2 = \|K\tilde{W}^l\|^2, \tag{B9}$$

it follows that

$$\begin{aligned}\sigma_l^l &\leqslant \Lambda^{1/2} \Gamma \, \|K\tilde{W}^l\| \\ &\leqslant \Lambda^{1/2} \Gamma \max_{A_l \in L_2} \min_{\substack{\tilde{W} \in A_l \\ \|\tilde{W}\| = 1}} \|K\tilde{W}\| \\ &\leqslant \Lambda^{1/2} \Gamma \mu_l. \end{aligned} \tag{B10}$$

We now estimate Λ and Γ. Recall that Λ is the largest eigenvalue of the Grammian (ψ_i^l, ψ_j^l). By Gershgorin's Theorem

$$|\Lambda - (\psi_i^l, \psi_i^l)| \leqslant \max_i \sum_{\substack{j=1 \\ j \neq i}} |(\psi_i^l, \psi_j^l)|. \tag{B11}$$

Observing that the ψ_i^l are normal and applying Schwarz's inequality we get

$$\Lambda \leqslant l. \tag{B12}$$

In a similar manner we obtain

$$\Gamma^2 \leqslant l. \tag{B13}$$

Thus from (B10)

$$\sigma_l^l \leqslant l\mu_l, \tag{B14}$$

a result already found (Corollary 3.4).

For the case of GSB the situation becomes more interesting. Assume Ψ_l is a GSB. Using (5.2) we find

$$(\psi_i^l, \psi_j^l) = 0, \qquad j \geqslant i + k, \tag{B15}$$

and (B11) yields

$$\Lambda \leqslant k. \tag{B16}$$

If P_l is also a GSB, then for some fixed k' dependent on P_l

$$\Gamma^2 \leqslant k'. \tag{B17}$$

We may state

THEOREM B1. *If Ψ_l and P_l are both* GSB *then the constant M''' in Theorem* 4.4 *may be replaced by M_l where $M_l \to \infty$.*

The proof is obvious upon noting Theorems 4.1 and 4.2.

APPENDIX C: GALERKIN MATRICES WITH CONDITION NUMBER UNITY

We shall prove in this appendix that for any kernel $K(x, y) \in L^2$ we may construct a set of Galerkin matrices all of which have $\gamma(\mathbf{K}^l) = 1$. This rather remarkable result suggests that the associated bases should be used in computations. It must be recalled, however, that success in numerical calculations depends not only upon conditioning but also upon truncation error. There is no assurance that the bases we are about to construct do a good job of approximating L_2. Thus the resulting truncation error may be totally unsatisfactory. Further investigation of this matter is called for.

For convenience in discussion we shall assume that $W_l = K(V_l)$. This simply avoids the need to use the projection operator Π_l introduced in Eq. (3.8). We take Ψ_l to be merely normal and proceed to find P_l such that $\mathbf{K}^l$ is a constant times a unitary matrix.

A preliminary result is needed. For convenience we shall say that a square matrix $\mathbf{A}$ has property (PR) if there exists a unitary $\mathbf{U}$ such that the matrix $\mathbf{Y}$ given by

$$\mathbf{Y} = \mathbf{A}\mathbf{U} \tag{C1}$$

is such that the diagonal elements of $\mathbf{Y}^*\mathbf{Y}$ are all equal.

THEOREM C1. *Every square matrix has property* (PR).

Proof. We write $\mathbf{A}$, assumed N by N, in terms of its singular value decomposition

$$\mathbf{A} = \mathbf{P}^*\mathbf{\Sigma}\mathbf{Q}, \tag{C2}$$

where $\mathbf{\Sigma}$ is the diagonal matrix of singular values and $\mathbf{R}$ and $\mathbf{Q}$ are the matrices of left and right singular vectors.

Let us assume that $\mathbf{\Sigma}$ has property (PR). Then we can find a unitary matrix $\mathbf{U}$ such that the diagonal elements of $(\mathbf{\Sigma}\mathbf{U})^*(\mathbf{\Sigma}\mathbf{U})$ are all equal.

Define

$$\mathbf{U}_1 = \mathbf{Q}^{-1}\mathbf{U}. \tag{C3}$$

Since $\mathbf{Q}$ is unitary, so is $\mathbf{U}_1$, and

$$\begin{aligned}(\mathbf{A}\mathbf{U}_1)^*(\mathbf{A}\mathbf{U}_1) &= \mathbf{U}_1^*\mathbf{A}^*\mathbf{A}\mathbf{U}_1\\ &= (\mathbf{Q}^{-1}\mathbf{U})^*\mathbf{A}^*\mathbf{A}(\mathbf{Q}^{-1}\mathbf{U})\\ &= (\mathbf{Q}^{-1}\mathbf{U})^*\mathbf{Q}^*\mathbf{\Sigma}\mathbf{P}\mathbf{P}^*\mathbf{\Sigma}\mathbf{Q}(\mathbf{Q}^{-1}\mathbf{U})\\ &= \mathbf{U}^*(\mathbf{Q}^*)^{-1}\mathbf{Q}^*\mathbf{\Sigma}\mathbf{P}\mathbf{P}^*\mathbf{\Sigma}\mathbf{Q}\mathbf{Q}^{-1}\mathbf{U}\\ &= (\mathbf{U}^*\mathbf{\Sigma})(\mathbf{\Sigma}\mathbf{U})\\ &= (\mathbf{\Sigma}\mathbf{U})^*(\mathbf{\Sigma}\mathbf{U}).\end{aligned} \tag{C4}$$

Since our assumption is that $\mathbf{\Sigma}$ has property (PR), clearly $\mathbf{A}$ does. Thus we may confine further investigation to diagonal matrices with non-negative elements.

Let x be a primitive Nth root of unity, $x^N = 1$. Define the N by N matrix:

$$\mathbf{E}^N = (x^{ij}). \tag{C5}$$

Since $|x^{ij}| = 1$, the norm or any row of column of $\mathbf{E}^n$ is $\sqrt{N}$. The (complex) scalar product of the ith row with the kth $(k \neq i)$ row is easily computed:

$$\sum_{j=1}^{N} x^{-ij}x^{kj} = \sum_{j=1}^{N} x^{-ij}x^{kj}$$
$$= \sum_{j=1}^{N} x^{(k-i)j}$$
$$= x^{k-i}\frac{1-x^{(k-i)N}}{1-x^{k-i}} = 0. \tag{C6}$$

Thus the matrix

$$\mathbf{F}^N = \frac{\mathbf{E}^N}{\sqrt{N}} \tag{C7}$$

is unitary.

Consider

$$\mathbf{Y} = \mathbf{\Sigma}\mathbf{F}^N, \tag{C8}$$

where $\mathbf{\Sigma}$ is the N by N diagonal matrix with entries σ_i. Note that

$$(\mathbf{\Sigma}\mathbf{F}^N)_{ij} = \frac{\sigma_i x^{ij}}{\sqrt{N}}. \tag{C9}$$

Hence all the diagonal elements of $\mathbf{Y}^*\mathbf{Y}$ are $(1/N)\sum_{j=1}^{N}\sigma_j^2$. Therefore any diagonal matrix has property (PR). (Observe that no use was made in the proof of the non-negativity of the σ_i.)

This establishes the theorem.

Remark. Note that the matrix $\mathbf{E}^N$ is really a Fourier transform matrix and is independent of $\mathbf{\Sigma}$. It can be shown that $\mathbf{E}^N$ is not unique, at least in the case that $N = 2^k$, for any k.

We now return to the principal matter at hand.

THEOREM C2. *Assume $W_l = K(V_l)$, with P_l any normal basis for W_l. Then there exists a normal basis Ψ_l for V_l such that the matrix $\mathbf{K}^l$ of K relative to these two bases is a constant times a unitary matrix. Moreover this constant is $l^{-1}\sum_{j=1}^{l}(\sigma_j^l)^{-2}$, where the σ_j^l are the singular values of a certain associated matrix, namely, the matrix of K relative to P_l and an orthonormal basis in V_l.*

Proof. Let Ξ_l form an orthonormal basis for V_l and let $\mathbf{K}_1^l$ be the matrix of K relative to the sets $\{\xi_i^l\}$ and $\{\rho_j^l\}$. From the previous theorem, $(\mathbf{K}_1^l)^{-1}$ has property (PR), so there exists a unitary matrix $\mathbf{U}$ such that if

$\mathbf{P}_1 = (\mathbf{K}_1^l)^{-1}\mathbf{U}$, then the diagonal elements of $\mathbf{P}_1^*\mathbf{P}_1$ are all equal to d^2 where

$$d^2 = \frac{1}{l} \sum_{j=1}^{l} \frac{1}{(\sigma_j^l)^2}, \tag{C10}$$

with σ_j^l the singular values of $\mathbf{K}_l$.

We now define a new set $\{\psi_k^l\}$ by

$$\psi_k^l = \frac{1}{d} \sum_{j=1}^{l} (\mathbf{P}_1 \mathbf{e}_k, \mathbf{e}_j)_l \, \xi_j^l. \tag{C11}$$

Here $\mathbf{e}_j$ is the usual unit basis vector with unity in the jth position. We claim that Ψ^l has the properties asserted in the theorem.

First, Ψ^l is a normal basis. For

$$\begin{aligned} \|\psi_k^l\|^2 &= \frac{1}{d^2} \left[\sum_{j=1}^{l} (\mathbf{P}_1 \mathbf{e}_k, \mathbf{e}_j)_l \, \xi_j^l, \sum_{j=1}^{l} (\mathbf{P}_1 \mathbf{e}_k, \mathbf{e}_i)_l \, \xi_i^l \right] \\ &= \frac{1}{d^2} \sum_{j=1}^{l} (\mathbf{P}_1)_{kj} (\mathbf{P}_1)_{kj} \\ &= \frac{1}{d^2} (\mathbf{P}_1^* \mathbf{P}_1)_{kk} = 1. \end{aligned} \tag{C12}$$

Now define

$$\begin{aligned} (\mathbf{K}^l)_{ij} &= (K\psi_i^l, \rho_j^l) \\ &= \frac{1}{d} \left[K \sum_{p=1}^{l} (\mathbf{P}_1 \mathbf{e}_i, \mathbf{e}_p)_l \, \xi_p^l, \rho_j^l \right] \\ &= \frac{1}{d} \sum_{p=1}^{l} (\mathbf{P}_1 \mathbf{e}_i, \mathbf{e}_p)_l (K\xi_p^l, \rho_j^l) \\ &= \frac{1}{d} \sum_{p=1}^{l} (\mathbf{P}_1)_{ip} (\mathbf{K}_1)_{pj} \\ &= \frac{1}{d} (\mathbf{P}_1 \mathbf{K}_1)_{ij} \\ &= \frac{1}{d} (\mathbf{U}^*)_{ij}. \end{aligned} \tag{C13}$$

Here we have used the definition of $\mathbf{P}_1$ at the last step. Since $\mathbf{U}^*$ is unitary, the theorem is established.

We may now state the principal result of this section.

THEOREM C3. *Assume $W_l = K(V_l)$, and let* P_l *be* any *normal basis for* W_l. *There always exists a normal basis* Ψ_l *for* V_l *such that the matrix* $\mathbf{K}^l$ *defined by* $(\mathbf{K}^l)_{ij} = (K\psi_i^l, \rho_j^l)$ *has condition number unity.*

Proof. The result is an immediate consequence of Theorem C2, since the singular values of $\mathbf{K}^l$ are all d^{-1}.

APPENDIX D: SOME FURTHER RESULTS ON SINGULAR VALUES

Recall the definition of μ_n^l (Eq. (3.5)) as the nth singular value of K restricted to V_l. Clearly, if V_l is the l-dimensional subspace of the first l singular functions of K, then $\mu_n^l = \mu_n$, $n = 1, 2, \ldots, l$. By choosing Ξ_l as any orthonormal subspace in V_l, we obtain from (C10)

$$d^2 = \frac{1}{l} \sum_{j=1}^{l} \frac{1}{(\mu_j^l)^2}. \tag{D1}$$

But d^{-1} is the (repeated) singular value of the matrix $\mathbf{K}^l$ found in Theorem C2, which we write as σ_l^l. Hence we have

THEOREM D1. *Let K be restricted to V^l. Then*

$$\frac{1}{(\sigma_l^l)^2} = \frac{1}{l} \sum_{j=1}^{l} \frac{1}{(\mu_j^l)^2}, \tag{D2}$$

where σ_l^l is the (repeated) singular value of the matrix representation of Theorem C2. If V_l is the space spanned by the first l singular values of K, then

$$\frac{1}{(\sigma_l^l)^2} = \frac{1}{l} \sum_{j=1}^{l} \frac{1}{(\mu_j)^2}. \tag{D3}$$

This result actually represents a sharpening of Theorem 3.2. While that theorem assumes the basis in $K(V_l)$ to be orthonormal and that in V_l to be merely normal, the roles may be switched by considering K^* instead of K. In fact, several of the results in this paper that involve one normal basis and one orthonormal have duals that may be obtained by replacing K by K^*.

We have one more observation:

THEOREM D2. *There exists a kernel K such that, in the notation of Theorem* D1,

$$\lim_{l \to \infty} \frac{(\sigma_l^l)^2}{l\mu_l^2} = 1. \tag{D4}$$

Proof. Define

$$K(x, y) = \sum_{j=1}^{\infty} \frac{1}{j!} \sin j\pi x \sin j\pi y, \qquad 0 \leqslant x,\ y \leqslant 1. \tag{D5}$$

It is easy to see that for this K,

$$\mu_l = \frac{1}{l!}. \tag{D6}$$

From (D3)

$$\frac{1}{(\sigma_l^l)^2} = \frac{1}{l} \sum_{j=1}^{l} (j!)^2. \tag{D7}$$

Hence

$$\begin{aligned} \frac{l\mu_l^2}{(\sigma_l^l)^2} &= \sum_{j=1}^{l} \left[\frac{j!}{l!}\right]^2 \\ &= 1 + \frac{1}{l^2} + \frac{1}{l^2(l-1)^2} + \cdots + \frac{1}{(l!)^2} \\ &= 1 + O\left(\frac{1}{l}\right) \end{aligned} \tag{D8}$$

and the result follows.

REFERENCES

1. G. M. WING, "Condition Numbers of Matrices Arising from the Numerical Solution of Linear Equations of the First Kind," Los Alamos National Laboratory Report LA-UR-82-3087, *Journal of Integral Equations*, in press.
2. R. C. ALLEN, JR., W. R. BOLAND, G. M. WING, Numerical experiments involving Galerkin and collocation methods for linear integral equations of the first kind, *J. Comput. Phys.* **49** (1983), 465–477.
3. J. HERSCH, Caractérisation variationnelle d'une somme de valeurs proper consécutives; généralisation d'inégalités de Pólya–Schiffer et de Weyl, *C. R. Acad. Sci.* **254** (1961), 1714–1716.

4. S. H. CHANG, A generalization of a theorem of Hille and Tamarkin with applications, *Proc. London Math. Soc.* (*2*) **3** (1952), 22–29.
5. E. HILLE AND J. D. TAMARKIN, On the characteristic values of linear integral equations, *Acta Math.* **57** (1931), 1–76.
6. C. DE BOOR, "A Practical Guide to Splines," Springer-Verlag, New York, 1978.
7. SLATEC Collection, Version 1.0, NESC No. 820, National Energy Software Center, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Ill. 60439.
8. J. J. DONGARRA, J. R. BUNCH, C. B. MOLER AND G. W. STEWART, "LINPACK User's Guide," SIAM, Philadelphia, 1979.
9. K. FAN, On a theorem of Weyl concerning eigenvalues of linear transformations, *Proc. Nat. Acad. Sci. U.S.A.* **35** (1949), 652–655.